United Nations E/ICEF/2011/15

Economic and Social Council

Distr.: General
18 July 2011

Original: English

For action

United Nations Children's Fund
Executive Board
Second regular session 2011
12-15 September 2011
Item 5 of the provisional agenda*

Annual report on the evaluation function and major evaluations in UNICEF

Summary

This annual report has been prepared in accordance with the Evaluation Policy (E/ICEF/2008/4) and relevant Executive Board decisions (2008/4; 2008/22; 2009/18; 2010/16). The report provides information on global developments in evaluation, the current state of the evaluation function in UNICEF at country, regional and global levels and outlines progress in strengthening the decentralized evaluation function. It also contains a summary of selected major evaluations conducted at country, regional and global levels within focus area 2 (basic education and gender equality) of the medium-term strategic plan. A draft decision is included in final section of this report.

* E/ICEF/2011/13.

11-41965 (E) 280711

Please recycle

Introduction

1. The evaluation function helps to ensure that UNICEF has timely, strategically focused and objective information on the performance of policies, programmes and initiatives to produce better results for children and women. It contributes to both organizational learning and accountability. The UNICEF Evaluation Office provides leadership for the evaluation function throughout the organization. In addition, it commissions independent evaluations and undertakes inter-agency evaluations within the United Nations system and joint evaluations with other partners.

2. This report is structured in eight parts. Section I describes major developments in United Nations system-wide evaluation coherence. Section II presents data and highlights from representative global evaluations. Section III presents actions and progress in strengthening the evaluation function, especially at a decentralized level. Section IV presents data from key performance indicators of the evaluation function. Section V describes UNICEF action in support of national evaluation capacity-building objectives. Section VI outlines the Evaluation Office workplan for 2012-2013. Section VII analyses the role of the evaluation function within the medium-term strategic plan (MTSP) focus area 2: basic education and gender equality. Section VIII contains a draft decision for consideration by the Executive Board.

I. System-wide evaluation coherence

3. UNICEF has an established history of engagement and leadership to strengthen accountability and coherence of the United Nations and international evaluation system. In recent years, the United Nations Evaluation Group (UNEG) has been the chief mode of managing this focus. UNICEF has maintained an instrumental role at UNEG, both in terms of advocacy for a strong evaluation function across the United Nations system and in leading UNEG's substantive work in several areas. At the decentralized level, regional offices and country offices have been working together with other United Nations agencies: the Asia-Pacific Shared Services Centre chairs the United Nations Development Evaluation Group for Asia and the Pacific; a number of regions have joint initiatives with other United Nations agencies; and many country offices are supporting United Nations Development Assistance Framework monitoring and evaluation task forces.

4. Over the past two years, UNICEF has had substantive involvement with areas of UNEG's work:

(a) The finalization and pilot testing of a guidance document on integrating gender equality and human rights in evaluation: this is developing core guidance to replace widely varied organizational approaches;

(b) Promotion of impact evaluation and development of guidance on impact evaluation issues: this helps agencies choose among statistically driven and mixed method approaches;

(c) Guidance on United Nations Development Assistance Framework (UNDAF) evaluations: this answers field demand for technical guidance to meet the evaluation accountability;

(d) Formulation of a concept paper for United Nations system collaboration on national capacity development in evaluation: this sets the stage for collaborative United Nations work, as discussed in Section V;

(e) Refining of norms and standards for evaluation function in the United Nations system: this updates the first UNEG product that helped ensure a consistent policy basis across disparate agencies.

5. UNEG has proven a capable forum for establishing minimum standards and basic guidance. It has increasingly shown strength in leading or supporting complex systemic evaluations. It is now ready to undertake joint activities on national capacity development, and to address a wider range of evaluation issues, including evaluation of the United Nations normative role as well as its humanitarian work. However, UNEG is constrained by the weak evaluation capacity of many United Nations agencies and by the uneven monitoring and evaluation capacity of country teams.

II. The UNICEF corporate evaluation agenda

A. The integrated monitoring and evaluation framework of the medium-term strategic plan

6. The integrated monitoring and evaluation framework (IMEF) contains commitments by UNICEF to conduct 15 corporate-level evaluations in 2010-2011. A resume of the IMEF evaluation results and management responses is available on the Executive Board website.[1]

7. The summary status of progress in implementing IMEF evaluations at end May 2011 was: completed: 7 of 15; under-implementation: 5; mobilization: 1; pending (no action): 2. This result is considered a satisfactory implementation rate at this point in the biennium since some commitments normally carry over into year 3. Some have had the focus or rigor level adjusted. Those still pending are typically waiting for necessary resources such as partner presence or key staff. Only one is considered unlikely to begin this biennium. The details are given in annex 1.

8. The individual IMEF evaluations are not summarized here. Two are discussed in section 7, and all completed ones are available online. More generally, this set of evaluations links with corporate learning priorities in the following ways:

(a) It includes global evaluations in each MTSP focus area and also in cross-cutting issues;

(b) It builds on the national and global impact measurement work of the UNICEF statistics section by exploring the causal relationships driving the impacts seen;

(c) Each evaluation examines major institutional effectiveness issues, to help UNICEF become more efficient and effective;

[1] http://www.unicef.org/about/execboard/index.html.

(d) Fostering effective partnership is a frequent key theme, including partnerships with national committees, other United Nations agencies, donor partners, and national stakeholders;

(e) Four of the 15 evaluations were conducted jointly with other agencies;

(f) At least 5 of the evaluations examined UNICEF work at the community level while others examined the work at upstream policy or global coordination levels;

(g) Five of the evaluations are centred on humanitarian issues and others include some humanitarian aspects;

(h) Although the agenda was developed before the equity re-focus emerged, virtually all the evaluations examine actions to reach the unreached or the underlying rights-based strategies and policies that guide corporate directions.

B. Inter-agency and multi-partner evaluations: "Delivering as One"

9. The country-led evaluation of seven of the "Delivering as One" pilot countries[2] was completed in late 2010 and provided valuable information on progress, challenges and lessons learned. Each country established and managed an autonomous evaluation of the five "Delivering as One" pillars (one plan, one leader, one house, one budgetary framework and one voice). UNICEF co-chaired the UNEG task force offering quality assurance. The country-led evaluations are being supplemented by a formal independent evaluation of the "Delivering as One" experience that is currently under way. Through UNEG, UNICEF is also contributing technical expertise to the independent evaluation.

10. The "Delivering as One" pilot findings are available through the Executive Board website. In summary, the findings suggest that United Nations Coherence brought some improvements to the work of the United Nations. There is a greater degree of collaboration, a more unified interface with the Governments, a stronger mechanism to mobilize core resources, and some evidence of costs saved. However, "Delivering as One" also poses challenges. Agencies are required to engage in One United Nations processes without a commensurate reduction in agency-specific planning, review and reporting requirements. The process may lead to fear and uncertainty about the implication for staffing. Some United Nations staff feel that "Delivering as One" has added an additional layer of bureaucracy and transaction costs.

C. Inter-agency and multi-partner evaluations: humanitarian response

11. UNICEF and other agencies supported an Evaluation of the Common Humanitarian Funds, led by the Office of the Coordinator for Humanitarian Affairs (OCHA). This was a $1.3 billion effort to support responses to long-term emergencies in 3 countries. The evaluation, completed in early 2011, reviewed the processes, outcomes, operational effects and operational impact of the Common

[2] Albania, Cape Verde, Mozambique, Rwanda, United Republic of Tanzania, Uruguay and Viet Nam.

Humanitarian Funds. OCHA and the partners are formulating a management response.

12. UNICEF, together with OCHA and the International Rescue Committee (representing all non-governmental organization [NGO] participants), co-managed an Inter-Agency Real-Time Evaluation of the response to the January 2010 earthquake in Haiti. A follow-up evaluation, focusing on issues of transition and national capacity development, is currently under way. The main UNICEF contribution to this second Inter-Agency Real-Time Evaluation is an inter-agency survey of the earthquake-affected population in Haiti based on a nationally representative sample, which UNICEF is managing on behalf of an inter-agency survey management group. This survey will yield vital impact data and also meet commitments to strengthening accountability to the affected population.

13. A review of the global education cluster co-leadership arrangement between UNICEF and Save the Children is discussed in section VII.

III. Strengthening evaluation within UNICEF

Context and overall strategy

14. Over 95 per cent of UNICEF-supported evaluations are authorized and managed in field offices, reflecting the decentralized nature of the organization. This ensures that evaluations generate contextually valid evidence, which is most likely to inform national policies for children. However, it also poses a managerial challenge to ensure that the field offices with their limited technical evaluation capacities actually produce good quality evaluations.

15. The approach that has been agreed upon focuses on support to field offices from the Evaluation Office and from regional offices. These two levels are working jointly to strengthen the evaluation function in a number of areas: strengthening governance; enhancing strategic planning of evaluations; promoting and supporting the quality of evaluations; improving the use of evaluations and management responses; and strengthening internal evaluation capacity. A related initiative reported in the next section is the development of national evaluation capacity both to support the monitoring and evaluation needs of the UNICEF supported programme and to ensure adequate evaluation skills for a broad range of public and private-sector needs.

16. This section highlights new information. Many initiatives reported in prior Executive Board reports remain operational but are not revisited here.

Recent developments in support of the strategy

A. Leadership and governance

17. In compliance with Executive Board decision 2010/16, the Director post was recruited at the D-2 level. The selected candidate has an extensive management history in evaluation, and was most recently the Director of Evaluation at the African Development Bank. The Director reports to the Office of the Executive Director through the Deputy Executive Director (Management), and is supported by

the Global Evaluation Committee, composed of the Executive Director, other senior leaders, regional directors, and an external expert. The Committee met in June 2010 and February 2011, and is scheduled to meet again in September 2011.

18. At the regional level, senior leadership improvements were seen in Asia, where each regional office will have a dedicated evaluation officer (at the P-5 level), compared to the shared post at present. Regional evaluation committees composed of representatives and regional office leaders have been created in 4 of the 7 regions. These committees advise on regional evaluation priorities and on improving adherence to the evaluation policy. Of special note, the Regional Office for the Americas and the Caribbean has authorized a full review of the monitoring and evaluation function in the region, including post design and post levels, independence, staff skills, links with counterparts. By the end of 2012, a variety of recommendations will be presented that may bring major shifts in staffing patterns and responsibilities and the use of external centres of excellence. The Evaluation Office will support this review, as needed, and other regions will look for relevant lessons when planning their 2014-2015 efforts.

19. At the country office level, there has been a steady increase in linking the evaluation function directly with senior management. In the Americas and the Caribbean, for example, 21 out of 24 country offices have the monitoring and evaluation specialist reporting to the Representative or Deputy Representative. In the West and Central Africa region, 9 out of 24 country offices have made the country management team led by the Representative responsible for the strategic selection of evaluations, rather than individual programme officers. However, there are also trends related to the monitoring and evaluation specialist function that bear watching. First, offices are increasingly combining the monitoring and evaluation function with other programme functions (e.g. social policy, knowledge management), which overloads monitoring and evaluation posts with too many responsibilities. Second, in some cases, the funding source is shifting from the support budget to other resources, without the funds in place to ensure that the post is filled. While these changes are allowed under the present policy, they clearly weaken the function. Ascertaining how widespread these changes are and preparing guidance on the issue will be a near-term priority.

B. Enhancing strategic planning of evaluations

20. As noted in the 2010 evaluation report, UNICEF issued a guidance note that described specific criteria for identifying priority topics for evaluation. Clearer guidance on the definition of evaluation was also developed, enabling UNICEF offices to better distinguish evaluations from surveys, studies and research. Substantial efforts to promote this guidance have since been made by all regional offices.

21. UNICEF presently emphasizes two mutually reinforcing directives: (a) that every programme component in a country programme be the subject of at least one evaluation per country programme cycle, in compliance with the evaluation policy; and (b) that each evaluation topic be selected for its importance according to the prioritization criteria and be adequately resourced.

C. Promoting and supporting the quality of evaluations

22. Independent assessments of the quality of the evaluation work done can help guide offices and senior management toward improvements. The 2010 evaluation report described the establishment of a Global Evaluation Reports Oversight System (GEROS), in which an external independent company had been hired to rate UNICEF-managed evaluations against UNEG evaluation standards. The four objectives of GEROS and the present status are as follows:

(a) Provide senior managers with an independent assessment of the quality of individual evaluation reports: successfully under way, all individual evaluations have a specific assessment;

(b) Report on the quality of evaluations reports commissioned by UNICEF offices: a UNICEF-wide summary report was delivered in March 2011. A second report has already been commissioned to cover evaluations undertaken in 2010;

(c) Strengthen internal evaluation capacity by providing commissioning offices with feedback, including practical recommendations on how to improve future evaluations: now being done for every completed evaluation sent to headquarters;

(d) Contribute to corporate knowledge management and organizational learning by identifying evaluation reports of good quality to be used in meta-analyses: sector-specific meta-analyses and lessons learned exercises were supported in 2010-2011 in HIV-AIDS, education, and child protection.

23. GEROS is now considered a core element of the quality assurance process and will be continued. Key indicators will be integrated into the new information system (VISION) for monitoring and reporting purposes.

24. In addition, regional offices invest considerable efforts at all phases to improve quality. All regions have in place quality assurance systems which review the draft integrated monitoring and evaluation plans, draft terms of reference and draft evaluation reports. More engagement by regional sectoral advisors would add important support to these systems.

D. Strengthening internal evaluation capacity

25. As explained in earlier reports, multiple efforts to improve internal capacity are being undertaken at headquarters and the regional level. These can be briefly summarized as a combination of knowledge management efforts (e.g. web-enabled communities of practice, web seminars, a help desk function) and more conventional network meetings, study visits and training seminars.

26. In addition, three further initiatives are planned. Building on the success of the web-based functions, UNICEF is developing an evaluation capacity development initiative which includes an online self-learning induction programme for new staff, as well as a training programme focusing on core skills (developing terms of reference, managing evaluations, writing evaluation reports, and developing management responses). Second, more attention will be paid in field offices to staff recruitment, job definition and structure, as well as to governance of the evaluation function. Third, the Evaluation Office is leading efforts to define how evaluation

can support the equity approach. A guidance package is under development, which will be complemented by a virtual resource centre and a series of web seminars.

E. Enhancing access to evaluations

27. Coordinated efforts by the Evaluation Office and Regional Offices have increased the timely submission rate of evaluation reports to the global reports database, from about 20 per cent in 2009 to 76 per cent in 2010.

IV. Key performance indicators

28. The cumulative impact of the systemic strengthening work should be clear in the performance of the evaluation system. This section presents for the first time a consolidated set of key performance indicators based on recent experiments in measuring system performance.

Indicator 1: Types of evaluations conducted

29. This measures the relevance of the evaluations and the adherence to good design principles. It responds to Executive Board decision 2010/16 requesting UNICEF to report on the types of evaluations it conducts. A set of definitions were developed and, as part of the GEROS quality review, external consultants were asked to categorize the 96 reports under review. As this is a first effort, the findings are indicative and require cautious interpretation.

A. Relevance

30. The purpose of each evaluation was defined as the overarching goal stated in the evaluation report. The great majority of evaluations (88 per cent) verified the results of UNICEF-supported programmes versus just 11 per cent that examined policy results delinked from field-based programmes. As UNICEF moves "upstream" the policy numbers should increase. Within the 88 per cent, 30 per cent examined pilot programming to determine suitability for scaling-up versus 55 per cent examining the results of a programme at its maximum intended scale. This is a healthy ratio, as it shows continued programme innovation as well as priority attention to the more numerous "at-scale" programmes.

31. The programme timing data showed that 44 per cent were formative evaluations designed for mid-course corrections versus 56 per cent investigating the results at the end of the programming period. This is a good balance of managerial and accountability intents.

B. Scope

32. More problematic findings emerge regarding correspondence with the MTSP. Fully 85 per cent of evaluations examined issues within just one MTSP focus area versus 15 per cent that examined two or more focus areas. This probably reflects the prevailing strong sectoral approach in UNICEF-supported programmes. Within the

sectoral emphasis, 25 per cent examined cross-cutting issues (gender, human rights-based approach, knowledge management, communication for development) as a major focus. A much smaller set (5 per cent) specifically examined institutional effectiveness in programming — issues of supply, finance, human resources and planning. Efficiency concerns mean that institutional effectiveness analyses should be more prominent.

33. Regarding the geographical dimension, fully 95 per cent of the evaluations had a national or subnational frame of reference, while 5 per cent were multi-country, regional or global. The prevalence of cross-border concerns and the utility of comparing national experiences would argue for more evaluations at larger scales. Of those with subnational and national frames of reference, the percentages for each category were almost exactly the same. This fits well with the equity agenda that examines the specific problems of lagging groups (subnational) while also pursuing results across the entire population (national).

C. Independence and management

34. It is assumed that the level of independence correlates with quality. Independent external evaluations assign control to evaluation professionals that are not at all affiliated with those being evaluated. This good practice occurred 46 per cent of the time. In contrast, independent internal evaluations (43 per cent) are managed by the division being evaluated using external consultants. Though more open to bias, the quality results were the same as the independent external evaluations. Only 5 per cent were self-evaluations by persons or units examining their own work and these were the poorest quality by far.

35. Management arrangements can also influence quality. UNICEF consulted with others but was responsible for all aspects of the evaluation in 58 per cent of the cases. In just 1 per cent of cases, the work was co-managed with other United Nations agencies and in 17 per cent with other non-United Nations agencies. In contrast with the findings of prior meta-evaluations, joint efforts were of poorer quality on average. The next report will verify whether this is a trend to investigate. Some 7 per cent were country-led, which could be seen as a positive sign of national ownership. Quality was poorer, however, for the country-led and joint evaluations.

36. This initial review of the types of evaluations conducted does not reveal major problems. However, UNICEF should also reflect on what the desired ratios should be, with a view to assessing where there is underinvestment or overinvestment.

Indicator 2: Quality of UNICEF evaluations

37. The GEROS quality review[3] discussed earlier examined a portfolio of 96 evaluations conducted by offices at all levels. Almost all were completed in 2009, so the results show the impact of quality improvement strategies of approximately 2004-2009.

[3] UNICEF Global Evaluation Reports Oversight System: Quality Review of Evaluation Reports, IOD PARC, March 2011.

38. The essential findings are the following, with comparisons noted with the earlier 2002 meta-evaluation:

(a) The percentage of evaluation reports rated as "good" quality rose from 20 per cent in 2002 to 36 per cent in 2010;

(b) Reports rated as "average" or "adequate" quality were stable, at 47 per cent in 2002 and 49 per cent in 2010;

(c) The percentage of poor quality reports fell from 33 per cent (2002) to 15 per cent (2010).

39. These findings are largely consistent with the quality trends detected and reported to the Executive Board in the reports of 2006, 2008, and 2009. UNICEF is confident that these are real improvement trends.

40. The most important substantive finding of the GEROS review was the following: "The central issue appears to be that [UNICEF] evaluators are far clearer about the theory of evaluation (purpose, objectives, methodology, data collection) than the processing and analysis of data that is generated". The positive conclusion is that management is using evaluations to ask important questions and defining what is needed to answer them. Weaknesses are more generally found in evaluation implementation, in deriving meaningful results from the collected data, and in the uptake of lessons and recommendations.

41. The quality review recommended that UNICEF prioritize building skills in developing high-quality terms of reference; develop and communicate an equity and human rights strategy for the evaluation function; invest in developing results frameworks for challenging thematic areas; and deliver extra support to country-led evaluations. A joint management response by headquarters and regional offices has been prepared.

Indicator 3: Management response

42. Executive Board decision 2008/4 mandated that all UNICEF evaluations have a management response. As described in the 2010 evaluation report, new guidance on management responses was developed and publicized, and a data base to receive the responses was developed. Submitting offices are also obligated to periodically update the actions taken versus the commitments made.

43. Intense advocacy and oversight has caused a significant jump, to 47 per cent, in completed responses for evaluations completed in 2010 and submitted for quality review versus just 10 per cent before the guidance went into effect. Regional adherence varies widely, falling as low as 20 per cent.

44. The rapid growth in compliance is encouraging. Further improvements are expected from two additional efforts: (a) quick feedback from the GEROS team to originating offices that can help to focus the management response; and (b) reminder messages to regional offices from senior management.

V. National evaluation capacity development

45. UNICEF has asked its field offices to intensify support to national evaluation capacity development, following recent Executive Board decisions. Many different efforts were discussed in prior reports and are not reiterated here. The major new efforts and results achieved in 2010 are described below.

A. Guidance and policy

46. In 2011, the UNEG Annual General Assembly endorsed the concept paper on "Possible roles for UNEG's members in national evaluation capacity development", developed by a UNEG task force co-led by UNICEF and the United Nations Development Programme (UNDP). The guidance focuses on (a) acting as a "knowledge broker" facilitating "South-South" generation and sharing of good practices on national evaluation systems; (b) promoting country-led evaluations and national evaluation systems, and (c) promoting the professionalization of evaluation and use of national capacity.

47. A newly written internal framework, Evidence for Children: Developing National Capacities for Country-led Evaluation Systems, gives the rationale for engaging in national evaluation capacity development. It outlines actions to strengthen the enabling environment, develop institutional and individual capacities and foster cooperation among stakeholders. The guidance supports widespread activity in this area, as described below.

B. Global learning

48. To facilitate learning and knowledge sharing on country-led monitoring and evaluation systems, UNICEF and partners launched in 2010 an interactive Web 2.0 platform, *MyM&E*, to share knowledge on country-led monitoring and evaluation. The *MyM&E* system is open to everyone. Capacities include a reference centre with items contributed by members, including agencies like UNICEF, and 1 live web seminars on development evaluation issues. For example, 2,106 participants attended 14 web seminars offered by the major participating agencies (UNDP, the World Food Programme, UN-Women, the International Labour Organization, the Rockefeller Foundation and global evaluation associations).

49. UNICEF is increasingly supporting South-South learning. Three examples shall stand for the whole. The Asia-Pacific Shared Services Centre facilitated support to Maldives by Sri Lankan government officials in the conceptualization of a National Framework for monitoring and evaluation. The Central and Eastern Europe and the Commonwealth of Independent States Regional Office — in cooperation with the United Kingdom Department for International Development and the World Bank — facilitated an exchange of experience between the Governments of Tajikistan and Serbia, in order for the former to gather evidence and experience about inter-ministerial coordination, data flow mechanisms and the development of a poverty-reduction strategy monitoring and evaluation system. This contributed to the decision by the Tajikistan Committee on Statistics to make socio-economic data freely accessible, and the Government decreed the establishment of monitoring and evaluation departments in all ministries. In Jordan,

the Middle East and North Africa Regional Office facilitated sharing an Egyptian experience of developing a university-level evaluation diploma, and negotiations with Jordanian universities are ongoing to deliver a similar training programme there.

C. Strengthening country-led monitoring and evaluation systems

50. Mapping of national monitoring and evaluation capacity development initiatives implemented by UNICEF has shown that 120 UNICEF offices are strengthening the institutional capacity of Governments; 39 are strengthening an enabling national environment for evaluation; and 56 are developing capacities of government and partners' staff. The following examples serve to illustrate these activities.

51. In Kenya, UNICEF assisted the monitoring and evaluation Directorate of the Ministry of Planning, National Development and Vision 2030 in establishing the National Integrated Evaluation System as a key accountability framework. The system is now fully functional, producing periodic performance reporting, expenditure tracking surveys and sector reviews. In Uganda, UNICEF supported government efforts to further develop a nascent national monitoring and evaluation system. Recent efforts included conducting a diagnostic of the public sector monitoring and evaluation systems, focusing on a set of international benchmarks for public-sector results-based management. Following the study, a national monitoring and evaluation policy was issued by the Office of the Prime Minister.

52. Country-led systems extend beyond the public sector. UNICEF has supported the establishment and professionalization of national, regional and international evaluation associations for many years. This continues, with efforts to move beyond the meetings, training events, and networking that still predominate. A major breakthrough was made by the Morocco Evaluation Association, which has attained credibility such that it is now invited to speak on evaluation matters in the public sphere. The latest milestone was the invitation by the Government to help enhance public accountability and evidence-based policy-making by contributing to the revision of the Constitution.

VI. The UNICEF Evaluation Office workplan 2012-2013

53. Executive Board decisions 2008/4 and 2008/22 requested that the biennial workplan of the UNICEF Evaluation Office be presented to the Executive Board. The UNICEF Evaluation Office workplan 2012-2013 defines 3 major contributions to UNICEF strategic goals.

54. By providing global leadership in evaluation within UNICEF and the global developmental community, the Evaluation Office will elicit improvements in management engagement, investment and partnerships. This main internal objective is guiding the Office to fulfil its role within the renewed UNICEF focus on equity. Within inter-agency fora, emphases will include updating the norms and standards for the evaluation function, undertaking joint evaluations on critical common concerns like "Delivering as One", and activating common approaches to national evaluation capacity development and UNDAF evaluations.

55. By conducting independent corporate evaluations, the Office will provide credible evidence on the performance of UNICEF programmes, policies and strategies. Clusters of evaluations will be conducted in programmatic themes linked to the MTSP focus areas, in humanitarian programming, and in cross-cutting institutional effectiveness issues. The institutional effectiveness portfolio will incorporate as one part the tools and methods developed within the joint programme performance assessment pilot of the Evaluation Office and the Office of Internal Audit. The 2010-2011 evaluation agenda that had been presented to the Board[4] remains in effect, but will be updated later in 2011. It will be important to coordinate this with the new Office of Research to ensure complementarity.

56. The Office provides leadership for systemic strengthening efforts by key stakeholders at the decentralized and global levels to meet their evaluation accountabilities. The improvement now being observed throughout the organization is expected to continue. The strategies of headquarters and regional offices discussed in section 4 will continue and expand (web-based connectivity, support to national evaluation capacity strengthening). A major objective is strengthening staff competence through enhanced staff recruitment, skills assessment strategies, and expanded learning options. A special emphasis will be placed on humanitarian monitoring and evaluation competencies following from weaknesses seen in multiple evaluations and reviews. Particularly strong links with regional office actions are planned for all systemic strengthening efforts.

57. For 2012-2013, spending by the Evaluation Office is expected to be $13.7 million from all sources. This is zero growth from the present. The staff complement remains the same (seven professionals on the Biennial Support Budget).

VII. Evaluation in basic education and gender equality (BEGE)

58. Prior reports to the Executive Board have included a summary of various evaluations conducted by UNICEF offices in the preceding year. These summaries had neither thematic coherence nor cross-cutting analysis and did not attract the interest of Board members. This year's report highlights evaluations from a single MTSP focus area: basic education and gender equality (BEGE). Following a brief introduction presenting data on the numbers of BEGE evaluations and studies and on BEGE evaluation expenditures, the major portion of this section reviews exemplary evaluations in each key result area (KRA) to determine whether they are generating relevant and useful knowledge.[5]

Evaluations and expenditures in the BEGE focus area

59. The strategic intent of BEGE is "to ensure that governments, communities and parents acquire the capacities and support necessary to fulfil their obligation to ensure the right of all children to free, compulsory quality education in all contexts, including humanitarian, recovery and fragile situations". BEGE programme

[4] E/ICEF/2010/10/Annex II.

[5] A general update on UNICEF work in education was provided in the Annual Report of the Executive Director (E/ICEF/2011/9).

activities are organized to achieve results in four KRAs: (a) school readiness; (b) access to quality basic education; (c) improving education quality; (d) restoring education after emergencies.

60. Table 1 below indicates that an estimated 195 evaluations and 433 studies and surveys were carried out in the BEGE programme area from 2006 to 2010 across all levels of UNICEF. This underrepresents the total, as some correlated evaluation efforts in other sectors are not counted (e.g. school water, sanitation and hygiene [WASH] services).

Table 1: Number of BEGE evaluations by key result area and year

	Evaluations						*Studies*
BEGE Key result areas	*2006*	*2007*	*2008*	*2009*	*2010*	*Total*	*Total*
KRA 1 — School readiness	1	5	4	5	1	**16**	**27**
KRA 2 — Access to quality basic education	10	16	15	19	5	**65**	**194**
KRA 3 — Improving education quality	21	26	21	23	16	**107**	**205**
KRA 4 — Education after emergencies	1	1	1	1	3	**7**	**17**
Total number of evaluations/studies	**33**	**48**	**41**	**48**	**25**	**195**	**443**

61. The numbers basically align with programme spending. For example, KRA 3 consumes about 48 per cent of education funds and conducted 55 per cent of the evaluations. The small percentage of education in emergencies evaluations is not worrisome as the uncounted multisectoral humanitarian evaluations normally include education.

62. UNICEF spent about 21 per cent of its programme budget on BEGE in 2009-2010, or $1.3 billion. Table 2 indicates how much of this was spent on evaluations and on studies and surveys, and also presents these as a percentage of funds spent on programmes.

Table 2: BEGE programme and evaluation expenditures by region

	Combined 2009-2010 BEGE expenditures		
Region	*Programmes in US$ millions*	*Evaluations in US$ millions (percentage of programme expenditures)*	*Studies and surveys in US$ millions (percentage of programme expenditures)*
Eastern and Southern Africa	333.0	1.7 *(0.5)*	9.0 *(2.8)*
South Asia	264.1	0.6 *(0.2)*	4.9 *(1.9)*
West and Central Africa	243.8	1.2 *(0.5)*	11.5 *(4.7)*
Middle East and North Africa	205.7	0.2 *(0.2)*	4.6 *(3.6)*
East Asia and Pacific	185.7	0.2 *(0.1)*	2.3 *(1.4)*
The Americas and Caribbean	109.7	0.3 *(0.2)*	2.7 *(2.3)*
Central and Eastern Europe and the Commonwealth of Independent States	42.0	0.1 *(0.3)*	1.8 *(4.3)*
UNICEF headquarters	18.7	2.3 *(12.0)*	Not available

63. UNICEF has previously reported that an average of 0.33 per cent of programme funds is used for evaluations. Just 2 of the 7 regions reach this very low percentage in BEGE. By contrast, studies and surveys and related research were more than 10 times larger than the evaluation spending.

64. These data are not contextualized enough to permit firm conclusions, but the following propositions warrant further study. Firstly, based on the relatively high expenditures on studies and surveys, UNICEF is making the necessary investments to testing solutions before they are packaged into programmes. This will be co-examined with the new Office of Research. Secondly, BEGE may be conducting enough evaluations (there is no firm metric) but, given the fact that BEGE evaluation spending is lower than the UNICEF average, it is probably not spending enough on each one. Thirdly, the high expenditure on global evaluations conducted at the headquarters level reflects the level of effort required to prepare rigorous evaluation designs, the scale of multi-country case studies, and the costs of experienced international and national evaluators.

Generating useful knowledge: lessons from select findings

65. This section includes findings from both global and national level BEGE evaluations. The global evaluations fairly represent the full set conducted in 2008-2010. As more than 90 per cent of evaluations are managed by UNICEF country offices, those cited in this report are a small and perhaps unrepresentative sample. However, they illustrate the kind of evidence and lessons generated. All of them were rated as "good" by the external quality review, and illustrate how the evaluation function can fulfil the role asked of it for learning, management decision-taking, and accountability.

66. **Key results area 1: Improving children's developmental readiness and starting school on time**. This KRA includes two global evaluations listed in the MTSP IMEF. In 2009/2010 the education section implemented a multi-country pilot of *Getting Ready for School*, a school readiness approach using structured learning to prepare children to enter primary school, and to prepare schools to receive younger children. In the pilot of the child-to-child approach, some 9000 preschool-aged children in six countries (Bangladesh, China, Democratic Republic of the Congo, Ethiopia, Tajikistan and Yemen) attended learning/play sessions facilitated by 2,000 "young facilitators", mostly older siblings of one or more learners in the group. An evaluation of the first year of the pilot found that there was a significant positive impact on on-time enrolment in Grade 1 in countries facing significant challenges to achieving universal primary education. Only China, which already had very high rates, did not show a significant impact. Other positive programme effects included increased parental support and participation in children's learning, and positive attitudes by young facilitators towards their own learning. However, the programme did not register the expected gains in terms of pre-primary literacy and numeracy skills. The evaluation is currently following-up with a comparison of the Grade 1 learners who participated in programme and the control group to see if there are differences in completion, class participation and learning.

67. In the second global evaluation, a multi-country evaluation of the early childhood development (ECD) programme supported by the Government of Netherlands found that in all sampled countries (Cambodia, Ghana, Nepal and

Tanzania), access to and coverage of ECD or pre-primary services has increased significantly. In all cases, ECD programming supported by UNICEF was found to be aligned with priorities for pre-primary education as expressed in national policies or government development plans. The evaluation also found that early learning and development standards have either been endorsed as a vital tool to define and monitor quality of ECD services or developed and implemented in all countries.

68. The evaluation also revealed some shortfalls. While some guidance for budgeting costs for ECD was developed, serious knowledge gaps were identified on costing and financing of ECD services. Data and monitoring problems include an absence of data on children's developmental progress due to ECD participation (Ghana) and weak data systems for planning and managing ECD activities and services (Nepal and Tanzania). Evidence on the sustainability of ECD programmes showed that governments had taken policy decisions to mainstream ECD into formal education systems but were not preparing adequate plans or budgets for scaling up.

69. **Key results area 2: Reduce disparities and increase access and completion of quality basic education**. While many more countries have made commitments towards educating their citizens, differential access to education remains a pervasive issue in most. KRA 2 focuses on providing more equitable access to education, especially for girls, for ethnic, linguistic or cultural minorities, and for out-of-school youth. It also focuses on influencing quality by attempting to mitigate factors that cause children to drop out of school.

70. One highly rated country-level evaluation examined education projects in the Osterode refugee camp for Kosovars in Romania. It found improvements in perceptions about schooling and willingness of children to attend school, an improvement in learning performance due to homework tutoring, heightened parenting skills and practices, and relatively good knowledge on environmental health, focused on a serious lead poisoning threat. Other evaluations involving vulnerable populations occurred in Uganda (alternative primary schooling for children of the Karamoja nomadic minority), Romania (access for Roma minority children), and China (access by children of internal migrant workers).

71. The Armenian school system is introducing "inclusive education" for children with disabilities. An evaluation showed inclusive schools were successfully changing attitudes and increasing services for students with disabilities. However, schools were found to have a very limited selection of "inclusive" practices that did not reach all types of disabilities or all needs of disabled children. As part of the broader UNICEF engagement with disability issues following the approval of the global convention, it is expected that both preparatory studies and evaluations will increase sharply in coming years.

72. Multilateral instruments and partnerships are critical for KRA 2 success, including the Education for All Fast Track Initiative, the School Fees Abolition Initiative, United Nations Girls' Education Initiative (UNGEI), Out-of-School Children Initiative, and sector-wide financing mechanisms. UNICEF has led or participated in joint evaluations for these initiatives and entities. An evaluation presently under way is the formative evaluation of the UNGEI partnership. It is asking if the partnership model increases efficiency and effectiveness, and if countries taking advantage of UNGEI-generated knowledge arrive at better solutions for improving girls' education.

73. **Key results area 3: Support to improve education quality, school retention, completion, and achievement rates**. UNICEF education quality improvement work is organized around the child-friendly schools (CFS) model. CFS approaches are cross-sectional, promoting and integrating quality with equity in WASH facilities, school design, HIV prevention and life-skills education, school feeding, disaster-risk reduction, gender equity and alternative forms of discipline. CFS initiatives address both structural factors (availability of safe classroom spaces and absence of qualified teachers) and teaching methods (adjusting traditional methodologies that do not cater to students' individual learning styles) that impede access to quality education.

74. CFS forms a major education programme component in many countries, which results in many evaluations. CFS approaches have also been evaluated through an independent global evaluation featured in the 2010 Executive Board report, and its findings recur regularly in country-level evaluations. Ministries of education embrace CFS principles to varying degrees, in part due to the advocacy of UNICEF and partners. CFS approaches bring education reform efforts under one umbrella, provide governments with a relevant framework for improving education quality, effectively engage stakeholders at all levels, foster child-focused teaching and learning, and encourage educators to respond to the needs of the child. There is no consistent evidence of an impact on learning achievement, repetition or drop-out rates other than positive effects on learning achievement seen in Sri Lanka. Undertaking more impact evaluations is a priority need in coming years.

75. The evaluations identified that CFS programming has been less successful in providing healthy school environments (sanitary and safe latrines; potable water) and providing education to cultural minorities. Equity concerns are registered especially for students with disabilities, as school heads and teachers reported overwhelmingly that they are not equipped to meet the needs of children with learning or developmental disabilities. As follow-up to the global evaluation, Australian Aid International has supported the UNICEF Education Section in developing a teacher education module and resource kit and in undertaking three case studies to improve educational quality for children with disabilities. Management responses to country-level evaluations commit UNICEF and partners to corrective or new programming strategies on the same issues (Armenia and Thailand).

76. **Key result area 4: Restoring education in emergencies and post-crisis situations**. UNICEF was involved in humanitarian action and post-crisis recovery in 79 countries in 2009 and in 98 countries in 2010. Education programming featured in most of these. The largest single investment in the current MTSP (the Netherlands, complemented by the European Community) has been the Education in Emergencies and Post-Crisis Transition (EEPCT) programme to bolster the education sector in 39 counties and 7 regional offices against the shocks that conflict and natural disasters can inflict on children's access to high-quality education.

77. A global evaluation of the EEPCT programme in 2010 found that education quality in emergencies and post-crisis settings had improved, although there was a lack of consistent understanding of the program's objectives. Child-friendly and protective school settings were found where there had been few or none, and another programme brought remedial educational opportunities to older children

who had missed out on their education during the crisis. At the same time, the evaluation found that prevention and preparedness pilots had not been tracked adequately (and cannot inform future efforts) and that monitoring and evaluation as a whole were weak, impeding the generation of credible assessment and learning. The evaluation also prompted UNICEF to think critically about the sustainability of one of its key resilience-building initiatives for displaced children in the border areas of West Africa.

78. As EEPCT funding was the largest-ever contribution to a single UNICEF programme, the evaluation also shed light on the challenges of managing large-scale contributions. The evaluation shows where UNICEF needs to improve its systems for managing, disbursing, monitoring and reporting on large programmatic funds. The UNICEF education section and other key divisions are implementing a management response to take these fiduciary management recommendations on board, alongside those related to programming.

79. The Global Education Cluster, a mechanism for supporting sectoral coordination in emergencies, formed the focus of a major evaluative review to assess the cluster's co-management by UNICEF and Save the Children. The evaluation highlighted that organizations representing different sectors (the United Nations system and an NGO) heighten the complexity of existing coordination challenges even when there are mutually reinforcing programme objectives. Underpinning these challenges have been divergent expectations as to whether the relationship between UNICEF and Save the Children should be a purely transactional one, based on identifying tasks and dividing up responsibilities between them, or a partnership in which both not only share responsibilities, but also resources, risks and accountabilities. In this vein, the review points to a lack of parity in the past, in which UNICEF controlled resources while Save the Children carried out much of the day-to-day work of the cluster, coupled with a lack of transparency and inclusion in information-sharing on the part of UNICEF.

80. The review notes UNICEF action to actively redress a number of these concerns, however, and offers suggestions for consolidating this progress — for example, by slightly modifying the cluster's architecture for enhanced accountability and by identifying measures to assess co-leadership functionality. A management response has been prepared by the UNICEF Education Section and other divisions.

Conclusion: sufficiencies and gaps in the BEGE evaluations

81. Overall, the BEGE programme area has had a robust evaluation agenda. Independent evaluations managed by the Evaluation Office have been conducted for every KRA, covering the largest programmes, where the majority of the BEGE investment goes. There is strong evidence of efforts to learn from evaluations and to integrate learning generated through evaluations into programmes. The most recent example of this is the *Getting Ready for School* evaluation.

82. Sufficient evaluation evidence is available to permit some systemic conclusions. For instance, the CFS global evaluation and more than 36 country-level CFS evaluations in 2009-2010 provide conclusive evidence about the efficacy and utility of CFS approaches. Lessons identified have a coherent theme: successful gains in inclusiveness and equity come from a deliberate effort to prepare school

facilities, personnel, programmes and learners. A critical mass of evidence is also available on responding to emergencies and reconstructing educations systems.

83. These successes notwithstanding, a number of important BEGE programming themes are yet to be reflected in the global evaluation agenda. These include the added-value of UNICEF involvement in national education-sector planning and in sector-wide approaches to development assistance; the organization's contribution and effectiveness in education systems strengthening; and the extent to which UNICEF has been successful in leveraging the four major global partnerships in education programming. As is the case with other programme sectors, BEGE evaluations have not offered any systematic learning on programme efficiency, partly because of the additional resources and expertise that these types of investigations entail.

84. To improve the quality of education evaluations and increase the likelihood of utilization of evaluation results, it will be necessary to increase the level of resources allocated to each evaluation. UNICEF country offices should also consider pooling resources and conducting joint evaluations with other development partners (for example other members that contribute to sector-wide approaches) who are working on similar programmes to provide governments with evidence on the efficacy of competing programme solutions. Such in-depth sector-wide evaluations would also provide a better basis for assessing the contribution of UNICEF within the education sector, and for sharing across the sector lessons from the organization's experience.

VIII. Draft decision

85. UNICEF recommends that the Executive Board approve the following draft decision:

The Executive Board

1. *Takes note* of the annual report on the evaluation function and major evaluations in UNICEF;

2. *Reaffirms* the central role played by the evaluation function in UNICEF and the importance of the principles set out in its Evaluation Policy (E/ICEF/2008/4);

3. *Welcomes* the evidence presented in the report of continued strengthening of the evaluation function, especially at the decentralized level, which contributes towards ensuring accountability, transparency, effectiveness and further improvement in the work of UNICEF;

4. *Takes note* of the introduction of key performance indicators tracking the effectiveness of the evaluation system;

5. *Welcomes* the thematic presentation of recent evaluation results in the annual report;

6. *Notes and encourages* UNICEF to:

(a) Sustain its engagement on evaluation within the United Nations system and more widely within the field of international development, including evaluation

of the response to humanitarian crises, with a view to strengthening accountability, transparency, coherence, effectiveness, learning and improvement;

(b) Sustain efforts to further strengthen the skills, capacities and systems of national partners, giving increased attention to South-South learning;

7. *Requests* UNICEF to:

(a) Ensure that management responses are prepared to address all evaluation recommendations, and to ensure implementation of agreed actions set out in such management responses;

(b) Ensure that relevant evaluation results are systematically considered and used in preparing key policies, strategies and programmes.

Annex 1

Status of medium-term strategic plan (MTSP) evaluations and research planned and conducted, 2010-2011

Status definitions:

Pending:	Still considered relevant but no action has been taken
Mobilization:	Terms of reference finalized or under design; resources being gathered
Implementation:	Field work under way or complete; analysis in process
Completed:	Final report delivered; dissemination under way
Superseded:	Original emphasis has shifted to a modified, higher priority

Title	Status at June 2011	Comment
Programmatic evaluations. Determine if a programming strategy central to MTSP success has (a) an evidence basis of impact; or (b) is being implemented with high levels of effectiveness.		
1. Evaluation of the UNICEF-Government of Netherlands early childhood development programme (Focus area 1)	Completed	
2. Community-based management of severe acute malnutrition (Focus areas 1 and 3)	Implementation	
3. Community case management (pneumonia, malaria, diarrhoea) (Focus area 1)	Implementation	
4. Global evaluation of the life-skills education programme (Focus area 2)	Implementation	
5. Enhancing school readiness through child-to-child programmes (Focus area 2)	Implementation	
6. Education in Emergencies and Post-crisis Transition programme (Focus area 2)	Completed	
7. Effectiveness of the Inter-agency Standing Committee Cluster for Emergency Education (Focus area 2)	Completed	Evaluation was converted to a review of the cluster's co-leadership arrangement. The review was completed and recommendations made to enhance co-management by the co-leads.
8. Evaluation of fostering social and behavioural change related to HIV prevention in adolescents (Focus area 3)	Pending	
9. Evaluation of the Unite For Children, Unite Against AIDS Campaign (Focus area 3)	Completed	

10. Global evaluation of the monitoring and reporting mechanism on grave child rights violations in situations of armed conflict (Focus area 4)	Mobilization	Multi-agency effort. Converted to a less rigorous good practices/ lessons learned study
Corporate-level topical or operational effectiveness evaluations. Determine if (a) the cross-cutting programming theme is successful; or (b) the organization is internally efficient. These require a well-developed set of examples from lengthy corporate engagement.		
1. Evaluation of community-based child care centres (Focus area 2)	Pending	
2. Organizational performance in adolescent development and participation programming (Focus area 5)	Completed	
3. Organizational performance in human rights-based approach to programming. (Focus area 5)	Implementation	
4. Year 1 Response to the Haiti Earthquake Humanitarian Crisis	Completed	Independent Review of UNICEF's Operational Response to the 2010 Earthquake in Haiti.
5. Inter-agency Real-Time Evaluation of Humanitarian Response 2010. (1) Haiti	Completed	